## AMENDMENTS TO THE CLAIMS

Please amend the claims as indicated hereafter.

**Claims:**

1-75. (Canceled)

76. (Canceled)

77. (Currently Amended) The method of claim ~~76~~ 80, wherein storing information related to said visual scene in a memory of the STT includes storing information identifying a location of said visual scene in relation to a point in said video presentation other than a point corresponding to a beginning of an entirety of the video presentation.

78. (Currently Amended) The method of claim ~~76~~ 80, wherein the video presentation is a video-on-demand presentation, and wherein the server transmits the portion of said video presentation starting from said visual scene responsive to the second user input.

79. (Canceled)

80. (Currently Amended) ~~The method of claim 79,~~ A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

receiving via a tuner in the STT a video presentation provided by the server;

outputting by the STT at least a portion of the video presentation as a television signal;

receiving a first user input associated with a visual scene contained in the video presentation;

storing information related to said visual scene in a memory of the STT responsive to receiving the first user input;

outputting by the STT at least another portion of the video presentation as a television signal after the information has been stored in the memory of the STT;

receiving a second user input configured to request said visual scene in said video presentation after the STT has output the at least another portion of the video presentation;

outputting by the STT a television signal comprising a portion of said video presentation starting from a location corresponding to said visual scene responsive to the second user input, wherein the location corresponding to said visual scene is identified by the STT using the information related to said visual scene;

receiving a user input configured to assign a character sequence to said visual scene in said video presentation;

storing data corresponding to said character sequence in a memory of the STT responsive to receiving the user input configured to assign a character sequence; and

providing said character sequence simultaneously with an image corresponding to said visual scene responsive to subsequent user input;

wherein said user input configured to assign a character sequence is received while said video presentation is being presented to said user.

81. (Canceled)

82. (Currently Amended) The method of claim ~~79~~ 80, further comprising receiving a plurality of user inputs configured to assign a plurality of respective character sequences corresponding to a plurality of respective visual scenes that were bookmarked responsive to a plurality of respective user inputs~~, wherein the plurality of user inputs configured to assign the plurality of respective character sequences are received after the video presentation has been provided to the user~~.

83. (Currently Amended) The method of claim ~~76~~ 80, further comprising the step of:

receiving a user input configured to request information related to said visual scene in said video presentation; and

providing the requested information responsive to receiving the user input configured to request information.

84. (Currently Amended) The method of claim ~~76~~ 80, wherein the first user input associated with the visual scene is received while the video presentation is being output by the STT ~~in a normal playback mode~~, wherein outputting the video presentation by the STT is not interrupted responsive to the first user input.

85. (Currently Amended) The method of claim 84, further comprising outputting information confirming that the visual scene has been bookmarked, wherein the information overlays a minority portion of a television screen being used to display the video presentation.

86. (Currently Amended) The method of claim 85, wherein said information confirming that the visual scene has been bookmarked includes at least one of a banner and an icon.

87. (Currently Amended) The method of claim ~~76~~ 80, further comprising storing information related to said visual scene in a memory of the server responsive to receiving the first user input.

88. (Canceled)

89. (Currently Amended) The method of claim ~~76~~ 80, wherein said second user input corresponds to a thumbnail image corresponding to the visual scene.

90. (Currently Amended) The method of claim ~~76~~ 80, wherein said visual scene is associated with a bookmark list associated with a plurality of visual scenes associated with a plurality of respective user inputs.

91. (Currently Amended) The method of claim ~~76~~ 80, further comprising associating a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective users responsive to a plurality of respective user inputs.

92. (Currently Amended) The method of claim ~~76~~ 80, further comprising associating a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective video presentations responsive to a plurality of respective user inputs.

93. (Currently Amended) The method of claim ~~76~~ 80, further comprising:

after expiration of a rental access period corresponding to the video presentation, prompting said user to provide input indicating whether said information is to be deleted from the memory of the STT.

94. (Currently Amended) The method of claim ~~76~~ 80, further comprising:

storing an image corresponding to said visual scene in a memory of the STT responsive to receiving the first user input;

95. (Currently Amended) The method of claim ~~76~~ 80, wherein said second user input requesting said visual scene corresponds to a thumbnail image corresponding to the visual scene, said thumbnail image being simultaneously provided with a plurality of thumbnail images corresponding to a plurality of visual scenes in the video presentation.

96. (Currently Amended) A television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said STT comprising:

a tuner configured to receive a motion video presentation provided by the server;

a memory;

a processor that is programmed to enable the STT to:

output at least a portion of the motion video presentation as a television signal;

store information related to a visual scene contained in the motion video presentation in the memory responsive to the STT receiving a first user input associated with said visual scene;

output at least another portion of the motion video presentation as a television signal after the information has been stored in the memory;

output responsive to the STT receiving a second user input a television signal comprising a portion of said motion video presentation starting from a location corresponding to said visual scene;

receive a user input configured to assign a character sequence to said visual scene;

store data corresponding to said character sequence in the memory responsive to receiving user input configured to assign a character sequence while said motion video presentation is being presented to said user; and

provide said character sequence simultaneously with an image corresponding to said visual scene;

wherein the location corresponding to said visual scene is identified by the STT using the information related to said visual scene; and

wherein the television signal comprising the portion of said motion video presentation starting from a location corresponding to said visual scene is output after the at least another portion of the motion video presentation is output as a television signal.

97. (Previously Presented) The STT of claim 96, wherein said visual scene is associated with a bookmark list associated with a plurality of visual scenes corresponding to a plurality of respective user inputs.

98. (Previously Presented) The STT of claim 96, wherein the processor is programmed to associate a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective users responsive to a plurality of respective user inputs.

99. (Currently Amended) The STT of claim 96, wherein the processor is programmed to associate a plurality of visual scenes with a plurality of respective bookmark lists associated with a plurality of respective motion video presentations responsive to a plurality of respective user inputs.

100. (Previously Presented) The STT of claim 96, wherein the processor is configured to prompt said user to provide input indicating whether said data is to be deleted from the memory of the STT.

101. (Previously Presented) The STT of claim 96, wherein the processor is configured to enable the STT to store in the memory an image corresponding to said visual scene responsive to receiving the first user input.

102. (Currently Amended) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

providing a plurality of images corresponding to a plurality of locations in a motion video presentation, the motion video presentation being received by the STT from the server via the bi-directional communication network, wherein each of the plurality of locations is associated with a respective user input received by the STT; and

providing a plurality of names corresponding to the plurality of images, wherein each of the plurality of names was selected by a respective user input

received by the STT while the motion video presentation was being output by the STT, wherein each of the plurality of names comprises a character sequence.

103. (Currently Amended) The method of claim 102, wherein at least one of the plurality of locations was identified by a respective user input while the motion video presentation was being output by the STT ~~in a normal play mode~~.

104. (Currently Amended) The method of claim 102, wherein at least one of the plurality of locations was identified by a respective user input while the motion video presentation was not being output by the STT.

105. (Previously Presented) The method of claim 102, wherein at least one of the plurality of names was selected by a respective user input from a list of names corresponding to one of the plurality of images.

106. (Currently Amended) A television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said STT comprising:

a processor programmed to enable the STT to output a plurality of images and a plurality of corresponding names, the plurality of images corresponding to a plurality of locations in a motion video presentation, the motion video presentation being received by the STT from the server via the bi-directional communication network, wherein each of the plurality of locations was identified by a respective user input received by the STT, and wherein each of the plurality of names was selected by a respective user input received by the STT while the motion video presentation was being output by the STT, and wherein each of the plurality of names comprises a character sequence.

107. (Currently Amended) The STT of claim 106, wherein at least one of the plurality of locations was identified by a respective user input while the motion video presentation was being output by the STT ~~in a normal play mode~~.

108. (Currently Amended) The STT of claim 106, wherein at least one of the plurality of locations was identified by a respective user input while the motion video presentation was not being output by the STT.

109. (Previously Presented) The STT of claim 106, wherein at least one of the plurality of names was selected by a respective user input from a list of names corresponding to one of the plurality of images.

110. (Currently Amended) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

identifying by the STT a plurality of locations in a motion video presentation responsive to a plurality of respective user inputs, the motion video presentation being received by the STT from the server via the bi-directional communication network;

associating by the STT a plurality of respective names with the plurality of locations responsive to a plurality of respective user inputs received by the STT while the motion video presentation was being output by the STT, wherein each of the plurality of respective names comprises a character sequence, and wherein the plurality of respective names include a first name and a second name, and wherein the plurality of locations include a first location and a second location;

outputting by the STT a first television signal configured to encode the first name and an image corresponding to the first location;

outputting by the STT a second television signal responsive to user input received while the first television signal was being output by the STT, the second

television signal being configured to encode the second name and ~~an~~ a second image corresponding to the second location.

111. (Currently Amended) The method of claim 110, further comprising:

receiving a user input corresponding to the second image; and

providing a portion of the motion video presentation starting from a location corresponding to the second image, responsive to receiving the user input corresponding to the second image.

112. (Currently Amended) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising steps of:

identifying a plurality of locations in a motion video presentation responsive to a plurality of respective user inputs, the motion video presentation being received by the STT from the server via the bi-directional communication network;

associating a plurality of respective names with the plurality of locations responsive to a plurality of respective user inputs received by the STT while the motion video presentation was being output by the STT, wherein each of the plurality of respective names comprises a character sequence;

providing a list that includes the plurality of names;

receiving user input corresponding to one of the plurality of names included in the list; and

providing a portion of the motion video presentation starting from a location corresponding to said one of the plurality of names.

113. (Currently Amended) The method of claim 112, wherein at least one of the plurality of locations was identified by a respective user input while the motion video presentation was being output by the STT ~~in a normal play mode~~.

114. (Previously Presented) The method of claim 112, wherein at least one of the plurality of names was selected by a respective user input from a list of names provided by the STT.

115. (Currently Amended) A method implemented by a television set-top terminal (STT) coupled via a bi-directional communication network to a server located remotely from said STT, said method comprising:

receiving via a tuner in the STT a motion video presentation provided by the server;

outputting by the STT at least a portion of the motion video presentation as a television signal;

receiving a first user input associated with a visual scene contained in the motion video presentation;

storing information related to said visual scene in a memory of the STT responsive to receiving the first user input;

outputting by the STT at least another portion of the motion video presentation as a television signal after the information has been stored in the memory of the STT;

receiving a second user input configured to request said visual scene in said motion video presentation after the STT has output the at least another portion of the motion video presentation; and

outputting by the STT a television signal comprising a portion of said motion video presentation starting from a location corresponding to said visual scene responsive to the second user input, wherein the location corresponding to said visual scene is identified by the STT using the information related to said visual scene;

receiving user input configured to assign a character sequence to said visual scene in said motion video presentation, wherein said user input configured to assign a character sequence is received by the STT while said motion video presentation is being output by the STT;

storing data corresponding to said character sequence in a memory of the STT

responsive to receiving the user input configured to assign a character sequence;

providing said character sequence simultaneously with an image corresponding to said visual scene responsive to user input;

receiving a user input configured to request information related to said visual scene in said motion video presentation;

providing the requested information responsive to receiving the user input configured to request information;

outputting information confirming that the visual scene has been bookmarked;

wherein the information overlays a minority portion of a television screen being used to display the motion video presentation;

wherein said information confirming that the visual scene has been bookmarked includes at least one of a banner and an icon;

wherein the motion video presentation is a video-on-demand presentation;

wherein the server transmits the portion of said motion video presentation starting from said visual scene responsive to the second user input;

wherein the first user input associated with the visual scene is received while the motion video presentation is being output by the STT in a normal playback mode; and

wherein outputting the motion video presentation by the STT is not interrupted responsive to the first user input.

116. (New) A method implemented by a television set-top terminal (STT), said method comprising:

receiving by the STT a first user input, said first user input being configured to assign a character sequence to a visual scene in a motion video presentation, said user input being received by the STT while the STT is outputting said motion video presentation;

storing data corresponding to said character sequence in a memory of the STT responsive to receiving the first user input; and

providing by the STT said character sequence simultaneously with an image

corresponding to said visual scene responsive to receiving a second user input;

receiving by the STT a third user input, said third user input corresponding to said visual scene; and

outputting a portion of said motion video presentation starting substantially from said visual scene responsive to receiving said third user input.

117. (New) The method of claim 116, wherein the image corresponding to said visual scene is a still image.

118. (New) The method of claim 117, further comprising:

outputting by the STT a plurality of still images corresponding to a plurality of visual scenes to the television responsive to receiving the second user input; and outputting by the STT a plurality of character sequences corresponding to the plurality of visual scenes to the television responsive to receiving the second user input;

wherein the plurality of still images and the plurality of character sequences are simultaneously displayed by the television.

119. (New) The method of claim 118, wherein the second user input is received by the STT while the outputting of the motion video presentation is suspended by the STT.